

M. Rao

1652

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/463,705

DATE: 08/14/2000  
TIME: 15:56:17

Input Set : A:\146.1335.txt  
Output Set: N:\CRF3\08142000\I463705.raw

Does Not Comply  
Corrected Diskette Needed

3 <110> APPLICANT: Cortes, J  
4 Fromentin, C  
5 Gaisser, S  
6 Leadlay, P  
7 Mendez, C  
8 Michel, J-M  
9 Raynal, M-C  
10 Salah-Bey, K  
11 Salas, J  
13 <120> TITLE OF INVENTION: BIOSYNTHESIS GENES & TRANSFER OF 6-DESOXY-HEXOSES IN  
14 SACCHAROPOLYSPORA ERYTHRAEA AND IN STREPTOMYCES  
15 ANTIBIOTICUS AND THEIR USE  
17 <130> FILE REFERENCE: 146.1335  
19 <140> CURRENT APPLICATION NUMBER: 09/463705  
20 <141> CURRENT FILING DATE: 2000-02-23  
22 <150> PRIOR APPLICATION NUMBER: PCT/FR98/01593  
23 <151> PRIOR FILING DATE: 1998-07-21  
25 <150> PRIOR APPLICATION NUMBER: 98/07411  
26 <151> PRIOR FILING DATE: 1988-06-12  
28 <150> PRIOR APPLICATION NUMBER: 97/09458  
29 <151> PRIOR FILING DATE: 1997-07-25  
31 <160> NUMBER OF SEQ ID NOS: 61  
33 <170> SOFTWARE: PatentIn Ver. 2.1

RECEIVED  
AUG 18 2000  
TC 1600 MAIL ROOM

ERRORED SEQUENCES

2689 <210> SEQ ID NO: 21  
2690 <211> LENGTH: 246  
2691 <212> TYPE: PRT  
W--> 2692 <213> ORGANISM:  
2694 <400> SEQUENCE: 21  
2695 Met Arg Ala Asp Thr Glu Pro Thr Thr Gly Tyr Glu Asp Glu Phe Ala  
2696 1 5 10 15  
2698 Glu Ile Tyr Asp Ala Val Tyr Arg Gly Arg Gly Lys Asp Tyr Ala Gly  
2699 20 25 30  
2701 Glu Ala Lys Asp Val Ala Asp Leu Val Arg Asp Arg Val Pro Asp Ala  
2702 35 40 45  
2704 Ser Ser Leu Leu Asp Val Ala Cys Gly Thr Gly Ala His Leu Arg His  
2705 50 55 60  
2707 Phe Ala Thr Leu Phe Asp Asp Ala Arg Gly Leu Glu Leu Ser Ala Ser  
2708 65 70 75 80  
2710 Met Leu Asp Ile Ala Arg Ser Arg Met Pro Gly Val Pro Leu His Gln  
2711 85 90 95  
2713 Gly Asp Met Arg Ser Phe Asp Leu Gly Pro Arg Val Ser Ala Val Thr  
2714 100 105 110  
2716 Cys Met Phe Ser Ser Val Gly His Leu Ala Thr Thr Ala Glu Leu Asp

Response for <213> organism is mandatory,  
genus/species, unknown, or artificial sequence.

\*Please check entire  
sequence listing as  
this error is recurring.

RAW SEQUENCE LISTING  
 PATENT APPLICATION: US/09/463,705

DATE: 08/14/2000  
 TIME: 15:56:17

Input Set : A:\146.1335.txt  
 Output Set: N:\CRF3\08142000\I463705.raw

```

2717      115      120      125
2719 Ala Thr Leu Arg Cys Phe Ala Arg His Thr Arg Pro Gly Gly Val Ala
2720      130      135      140
2721 Ile Glu Pro Trp Trp Phe Pro Glu Thr Phe Thr Asp Gly Tyr Val
E--> 2722 145      150      155      160
2724 Ala Gly Asp Ile Val Arg Val Asp Gly Arg Thr Ile Ser Arg Val Ser
E--> 2725      165      170      175
2727 His Ser Val Arg Asp Gly Gly Ala Thr Arg Met Glu Ile His Tyr Val
E--> 2728      180      185      190
2730 Ile Ala Asp Ala Glu His Gly Pro Arg His Leu Val Glu His His Arg
E--> 2731      195      200      205
2733 Ile Thr Leu Phe Pro Arg His Ala Tyr Thr Ala Ala Tyr Glu Lys Ala
E--> 2734      210      215      220
2736 Gly Tyr Thr Val Glu Tyr Leu Asp Gly Gly Pro Ser Gly Arg Gly Leu
E--> 2737 225      230      235      240
2739 Phe Val Gly Thr Arg Thr
E--> 2740      245

```

*See attached  
pages.*

09/463,705

# Sequence #20

<400> 20

atg cgg gct gac acg gag ccg acc acc ggg tac gag gac gag ttc gcc	48
Met Arg Ala Asp Thr Glu Pro Thr Thr Gly Tyr Glu Asp Glu Phe Ala	
1 5 10 15	
 gag atc tac gac gcc gtg tac cgg ggc cgg ggc aag gac tac gcc ggc	96
Glu Ile Tyr Asp Ala Val Tyr Arg Gly Arg Gly Lys Asp Tyr Ala Gly	
20 25 30	
 gag gcg aag gac gtg gcg gac ctc gtg cgc gac cgg gtg ccg gac gcg	144
Glu Ala Lys Asp Val Ala Asp Leu Val Arg Asp Arg Val Pro Asp Ala	
35 40 45	
 tcc tcc ctc ctg gac gtg gcc tgc ggc acg ggc gcg cac ctg cgg cac	192
Ser Ser Leu Leu Asp Val Ala Cys Gly Thr Gly Ala His Leu Arg His	
50 55 60	
 ttc gcc acg ctc ttc gac gac gcc cgc ggt ctc gaa ctg tcc gcg agc	240
Phe Ala Thr Leu Phe Asp Asp Ala Arg Gly Leu Glu Leu Ser Ala Ser	
65 70 75 80	
 atg ctg gac atc gcc cgc tcc cgc atg ccg ggc gtg ccg ctg cac caa	288
Met Leu Asp Ile Ala Arg Ser Arg Met Pro Gly Val Pro Leu His Gln	
85 90 95	
 ggg gac atg cga tcc ttc gac ctg ggg cca cgc gtc tcc gcg gtc acc	336
Gly Asp Met Arg Ser Phe Asp Leu Gly Pro Arg Val Ser Ala Val Thr	
100 105 110	
 tgc atg ttc agc tcc gtc ggc cac ctg gcc acc acc gcc gaa ctc gac	384
Cys Met Phe Ser Ser Val Gly His Leu Ala Thr Thr Ala Glu Leu Asp	
115 120 125	
 gcg acg ctg cgg tgc ttc gcc cgg cac acc cgg ccc ggc ggc gtg gcc	432
Ala Thr Leu Arg Cys Phe Ala Arg His Thr Arg Pro Gly Gly Val Ala	
130 135 140	
 gtc atc gaa ccg tgg tgg ttc ccg gag acc ttc acc gac ggc tac gtg	480
Val Ile Glu Pro Trp Trp Phe Pro Glu Thr Phe Thr Asp Gly Tyr Val	
145 150 155 160	

→ In sequence #21 following, it appears that this particular Valine residue was somehow deleted, thus causing errors in the amino acid numbering region, and an invalid amino count. (245 instead of 246)

09/463, 705

<210> 21  
<211> 246  
<212> PRT  
<213>

# Sequence #21

<400> 21

Met Arg Ala Asp Thr Glu Pro Thr Thr Gly Tyr Glu Asp Glu Phe Ala  
1 5 10 15

Glu Ile Tyr Asp Ala Val Tyr Arg Gly Arg Gly Lys Asp Tyr Ala Gly  
20 25 30

Glu Ala Lys Asp Val Ala Asp Leu Val Arg Asp Arg Val Pro Asp Ala  
35 40 45

Ser Ser Leu Leu Asp Val Ala Cys Gly Thr Gly Ala His Leu Arg His  
50 55 60

Phe Ala Thr Leu Phe Asp Asp Ala Arg Gly Leu Glu Leu Ser Ala Ser  
65 70 75 80

Met Leu Asp Ile Ala Arg Ser Arg Met Pro Gly Val Pro Leu His Gln  
85 90 95

Gly Asp Met Arg Ser Phe Asp Leu Gly Pro Arg Val Ser Ala Val Thr  
100 105 110

Cys Met Phe Ser Ser Val Gly His Leu Ala Thr Thr Ala Glu Leu Asp  
115 120 125

Ala Thr Leu Arg Cys Phe Ala Arg His Thr Arg Pro Gly Gly Val Ala  
130 135 140

Ile Glu Pro Trp Trp Phe Pro Glu Thr Phe Thr Asp Gly Tyr Val  
145 150 155 160

Val

here would correct the error. No adjustment to  
the numbering sequence is required.

## VERIFICATION SUMMARY

PATENT APPLICATION: US/09/463,705

DATE: 08/14/2000

TIME: 15:56:18

Input Set : A:\146.1335.txt

Output Set: N:\CRF3\08142000\I463705.raw

L:42 M:351 W: Sequence data Name/Key Feature Out-of-Range, SEQ ID#:1, CDS LOCATION: Complement((48)..(1046))  
L:48 M:351 W: Sequence data Name/Key Feature Out-of-Range, SEQ ID#:1, CDS LOCATION: Complement((2322)..(3404))  
L:173 M:201 W: Mandatory field data missing, ORGANISM  
L:243 M:201 W: Mandatory field data missing, ORGANISM  
L:324 M:351 W: Sequence data Name/Key Feature Out-of-Range, SEQ ID#:4, CDS LOCATION: complement (4)..(1266)  
L:378 M:201 W: Mandatory field data missing, ORGANISM  
L:603 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:607 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:710 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:714 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:718 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:722 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:726 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:730 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:734 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:738 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:742 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:746 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:750 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:754 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:758 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:762 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:766 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:772 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:776 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:780 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:784 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:788 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:792 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:796 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:800 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:804 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:808 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:812 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:816 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:820 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:824 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:828 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:832 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:836 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:840 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:844 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:848 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:852 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:856 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:860 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:864 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:868 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6

## VERIFICATION SUMMARY

PATENT APPLICATION: US/09/463,705

DATE: 08/14/2000

TIME: 15:56:18

Input Set : A:\146.1335.txt

Output Set: N:\CRF3\08142000\I463705.raw

L:872 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:876 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:880 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:884 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:888 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:892 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:896 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:900 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:6  
L:1122 M:201 W: Mandatory field data missing, ORGANISM  
L:1191 M:201 W: Mandatory field data missing, ORGANISM  
L:1276 M:201 W: Mandatory field data missing, ORGANISM  
L:1328 M:201 W: Mandatory field data missing, ORGANISM  
L:1430 M:201 W: Mandatory field data missing, ORGANISM  
L:1530 M:201 W: Mandatory field data missing, ORGANISM  
L:1696 M:201 W: Mandatory field data missing, ORGANISM  
L:2269 M:201 W: Mandatory field data missing, ORGANISM  
L:2353 M:201 W: Mandatory field data missing, ORGANISM  
L:2441 M:201 W: Mandatory field data missing, ORGANISM  
L:2529 M:201 W: Mandatory field data missing, ORGANISM  
L:2692 M:201 W: Mandatory field data missing, ORGANISM  
L:2722 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:21  
M:332 Repeated in SeqNo=21  
L:2740 M:252 E: No. of Seq. differs, <211>LENGTH:Input:246 Found:245 SEQ:21  
L:2746 M:201 W: Mandatory field data missing, ORGANISM  
L:2758 M:201 W: Mandatory field data missing, ORGANISM  
L:2770 M:201 W: Mandatory field data missing, ORGANISM  
L:2782 M:201 W: Mandatory field data missing, ORGANISM  
L:2806 M:201 W: Mandatory field data missing, ORGANISM  
L:2818 M:201 W: Mandatory field data missing, ORGANISM  
L:2830 M:201 W: Mandatory field data missing, ORGANISM  
L:2842 M:201 W: Mandatory field data missing, ORGANISM  
L:2854 M:201 W: Mandatory field data missing, ORGANISM  
L:2866 M:201 W: Mandatory field data missing, ORGANISM  
L:2878 M:201 W: Mandatory field data missing, ORGANISM  
L:2890 M:201 W: Mandatory field data missing, ORGANISM  
L:2902 M:201 W: Mandatory field data missing, ORGANISM  
L:2914 M:201 W: Mandatory field data missing, ORGANISM  
L:2926 M:201 W: Mandatory field data missing, ORGANISM  
L:2938 M:201 W: Mandatory field data missing, ORGANISM  
L:2950 M:201 W: Mandatory field data missing, ORGANISM  
L:2962 M:201 W: Mandatory field data missing, ORGANISM  
L:2974 M:201 W: Mandatory field data missing, ORGANISM  
L:2986 M:201 W: Mandatory field data missing, ORGANISM  
L:2998 M:201 W: Mandatory field data missing, ORGANISM  
L:3010 M:201 W: Mandatory field data missing, ORGANISM  
L:3022 M:201 W: Mandatory field data missing, ORGANISM  
L:3036 M:201 W: Mandatory field data missing, ORGANISM  
L:3048 M:201 W: Mandatory field data missing, ORGANISM  
L:3060 M:201 W: Mandatory field data missing, ORGANISM

VERIFICATION SUMMARY                      DATE: 08/14/2000  
PATENT APPLICATION: US/09/463,705        TIME: 15:56:18

Input Set : A:\146.1335.txt  
Output Set: N:\CRF3\08142000\I463705.raw

L:3072 M:201 W: Mandatory field data missing, ORGANISM  
L:3086 M:201 W: Mandatory field data missing, ORGANISM  
L:3098 M:201 W: Mandatory field data missing, ORGANISM  
L:3110 M:201 W: Mandatory field data missing, ORGANISM  
L:3122 M:201 W: Mandatory field data missing, ORGANISM  
L:3134 M:201 W: Mandatory field data missing, ORGANISM  
L:3146 M:201 W: Mandatory field data missing, ORGANISM  
L:3158 M:201 W: Mandatory field data missing, ORGANISM  
L:3170 M:201 W: Mandatory field data missing, ORGANISM